

# Sujit Silas Armstrong Suthahar

sujitsilas@g.ucla.edu | LinkedIn | 951-756-0678 | 840 Hilgard Ave, Los Angeles, CA, 90024

## EDUCATION

<b>University of California, Los Angeles (UCLA), Los Angeles, CA</b>   M.S. Bioengineering	Expected, March, 2025
<ul style="list-style-type: none"><li>GPA: 3.97/4.00</li></ul>	
<b>University of California, Riverside (UCR), Riverside, CA</b>   B.S. Biology, Honors	2021
<ul style="list-style-type: none"><li>GPA: 3.50/4.00</li></ul>	

## SKILLS

**Programming:** Proficient in Python, R, Bash/Shell, Go, SQL, JavaScript, C/C++, CSS/HTML, and LaTeX  
**Software:** MS Word, PowerPoint, Excel, Pymol, Linux, Blender, AutoCAD, IBM SPSS, SAS/STATA, Docker, Git  
**Machine Learning / Deep Learning:** TensorFlow, Keras, PyTorch, Scikit-learn, CUDA, TensorLy, numpy, GLIPH (TCR analysis)  
**Bioinformatics Tools:** STAR aligner, Bowtie2, GATK4, SALMON, DESeq2, Seurat, Scanpy, NextFlow, PyMol, BioPython, Bioconductor, PSIPRED, DeepFold, edgeR, dplyr, tidyr, ggplot2, CibersortX, Cytoscape, matplotlib, seaborn, WGCNA, MiXCR, MethylKit, ChIPseeker, DSS, GSEA, GenomicRanges, Monocle3, VEP

## WORK EXPERIENCE

<b>UCLA Institute for Society and Genetics</b>	Los Angeles, CA
<i>Graduate Student Researcher (GSR)</i>	09/2023 - Present
<ul style="list-style-type: none"><li>Spearheaded mapping and analysis of RRBS, WGBS, ATAC-seq, and RNA-seq data reducing turnaround times by 25%; analyzed bisulfite sequencing data and RNA-seq data using matrix factorization methods identifying key regulative elements</li><li>Devised NextFlow pipelines for both local and distributed computing environments, enhancing reproducibility while reducing data preprocessing times by 30% of bioinformatics workflows</li><li>Created a user-friendly HTML-based custom bash script generator specifically for nf-core pipelines, empowering non-expert users to effortlessly submit and manage jobs on high-performance computing clusters</li></ul>	

## GRAIL

<i>Computational Biology &amp; Machine Learning Intern</i>	Menlo Park, CA
	06/2024 - 09/2024
<ul style="list-style-type: none"><li>Optimized classifier features and Galleri assay probes by implementing a novel false positive investigation toolkit increasing cell-free DNA detection sensitivity for a few early-stage cancer types by 8%</li><li>Deployed a genome annotation pipeline reducing non-specific probe selection by 10% through DMR analysis, resulting in a revamped targeted panel and improved classifier features</li><li>Refactored functions, updated pipelines, and landed two merge requests related to development of two vignettes and false positive analysis workflow</li></ul>	

## UCLA Biomedical AI Research Lab

<i>Researcher</i>	Los Angeles, CA
	09/2023 - 06/2024
<ul style="list-style-type: none"><li>Implemented an ensemble-based deep learning model using multiple instance learning, achieving 76% diagnostic accuracy for indeterminate thyroid nodules—a 25% improvement over traditional methods and decreasing unnecessary biopsies</li><li>Leveraged CUDA architecture to accelerate deep learning model training on GPUs within a high-performance computing (HPC) environment, cutting down processing time and improving model iteration speed</li><li>Developed an advanced U-Net architecture featuring enhanced skip connections and integrated image preprocessing techniques achieving a sensitivity of 60% in biopsy needle segmentation problems for medical imaging</li></ul>	

## La Jolla Institute for Immunology

<i>Bioinformatics Specialist</i>	San Diego, CA
	06/2021 - 07/2023
<ul style="list-style-type: none"><li>Engineered automated pipelines for multi-modal single-cell RNA-seq analysis (including CITE-seq, and TCR-seq), cutting down processing time by 40% and standardizing analysis across 200 PBMC samples through R/Python packages</li><li>Developed and applied a computational method to validate and optimize antibody titration concentrations for a panel of 124 antibodies used in CITE-seq experiments, improving accuracy and reliability of multi-omics data integration</li><li>Designed a statistical framework for analyzing 348,000+ TCR sequences from atherosclerotic patients, identifying 672 APOB-reactive T-cell clones and 532 conserved motifs, enabling first-ever tracking of disease-specific T-cells with 80% cross-donor validation</li><li>Applied CIBERSORTx deconvolution to quantify immune cell fractions in bulk RNA-seq PBMC samples of CVD patients, revealing T-regulatory cell depletion, and developed single cell enrichment analysis pipeline characterizing 5 novel cell types in scRNA-seq data</li></ul>	

## University of California, Riverside

<i>Research Assistant (RA)</i>	Riverside, CA
	10/2019 - 06/2021
<ul style="list-style-type: none"><li>Conducted in-depth mRNA transcriptome analysis using R and Python, spearheading an independent research project comparing hESC transcriptomes of differentiating neural crest and ectodermal cells</li><li>Employed advanced protein modeling tools including Pymol, I-Tasser, PSIPRED, Robetta, and SAVES 6.0 to study DNA-transcription factor (TF) interactions in early development, integrating molecular biology knowledge with computational approaches to gain novel insights</li><li>Utilized fluorescence microscopy techniques to examine chick embryos providing critical insights into early developmental processes through high resolution fluorescence imaging while driving down imaging times by 20% with a new imaging protocol</li></ul>	

## LEADERSHIP EXPERIENCE

### Danish Display

Tamil Nadu, India

#### Marketing Director

08/2020 - 01/2021

- Conceptualized and executed a brand re-design, elevating product, and store appeal, driving a 30% increase in revenue
- Promoted digital marketing efforts, including website optimization and social media engagement, achieving a 46% growth in followers, and revamping business visibility
- Led influencer marketing strategy across 10+ partnerships, leveraging data analytics to revamp campaign ROI, resulting in 25% revenue growth and 3X engagement rates within 4 months

### Highlander Union

Riverside, CA

#### HUB Associate

09/2018 - 07/2020

- Fostered student success by facilitating access to campus resources, exemplifying a student-centric approach to service
- Coordinated 100+ campus events and managed 50+ venue bookings per semester, accomplishing 95% user satisfaction and decreasing scheduling conflicts by 40% through streamlined digital booking system
- Partnered with professional staff members to launch 3 new student programs, increasing event participation and attaining 90% positive feedback from 300+ students

## SELECTED PUBLICATIONS

- Armstrong Suthahar SS, Nettersheim FS, Alimadadi A, Wang E, Billitti M, Resto-Trujillo N, Roy P, Hedrick CC, Ley K, Orecchioni M. Olfr2-positive macrophages originate from monocytes proliferate in situ and present a pro-inflammatory foamy-like phenotype. Cardiovasc Res. 2024 Sep 4;cvae153. doi: 10.1093/cvr/cvae153. Epub ahead of print. PMID: 39229899
- Athreya, S., Melehy, A., Armstrong Suthahar, S. S., Ivezi, V., Radhachandran, A., Sant, V., ... & Speier, W. (2024). Reducing Overtreatment of Indeterminate Thyroid Nodules Using a Multimodal Deep Learning Model. arXiv preprint arXiv:2409.19171
- Armstrong Suthahar, S. S., ROY, P., Makings, J., & Ley, K. (2023). Identification of apolipoprotein B-reactive CDR3 motifs allow tracking of atherosclerosis-related memory CD4+ T cells in multiple donors. Frontiers in Immunology, 15, 1302031
- Armstrong, S. S., Nettersheim, F. S., Durant, C., Blanco-Dominguez, R., Roy, P., Orecchioni, M., Suryawanshi, V., & Ley, K. Titration of 124 antibodies using CITE-Seq on human PBMCs. Scientific Reports, 12(1), Article 1 (2022)
- Vallejo, J., Saigusa, R., Gulati, R., Armstrong Suthahar, S. S., et al. Combined protein and transcript single-cell RNA sequencing in human peripheral blood mononuclear cells. BMC Biol 20, 193 (2022)

## PROJECTS

- Survival benefit of liver transplant in mechanically ventilated patients
- tcrDiff - R package for differential T-cell receptor motif analysis
- R Shiny app for scRNA-seq and RNA-seq data visualization
- Leveraged machine learning algorithms to predict heart failure based on clinical records
- U-Net based deep learning model for lung biopsy needle segmentation in CT-scans

## RELEVANT COURSEWORK

Computer Vision, Linear Algebra, Machine Learning and Data Driven Modelling, Signal and Image Processing, Data Structures in Medical Informatics, Natural Language Processing, Deterministic Modelling, Statistical Mechanics, Advanced Statical Methods, Bayesian Statistics

## AWARDS & PROFESSIONAL ASSOCIATIONS

UCLA Graduate Student Researcher, American Association of Immunologists Travel Award, Highlander Union Excellence Award UCR Achievement Scholarship, American Association of Immunologists - Member, Graduate Faculty Executive Commission UCLA - Samueli School of Engineering Representative, UCLA Venture Accelerator - Member